

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/530,983  
Source: PT  
Date Processed by STIC: 2/15/06

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 02/15/2006

PATENT APPLICATION: US/10/530,983

TIME: 10:09:23

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02152006\J530983.raw

4 &lt;110&gt; APPLICANT: Innovations Foundation

8 &lt;120&gt; TITLE OF INVENTION: Secreted Acid Phosphatase (sapM) is Present Only in Pathogenic

9 Mycobacteria and Expressed Selectively at Phagosomal pH

12 &lt;130&gt; FILE REFERENCE: 4146 0005

C--&gt; 14 &lt;140&gt; CURRENT APPLICATION NUMBER: US/10/530,983

C--&gt; 16 &lt;141&gt; CURRENT FILING DATE: 2005-04-11

18 &lt;150&gt; PRIOR APPLICATION NUMBER: US 60/416,957

20 &lt;151&gt; PRIOR FILING DATE: 2002-10-09

22 &lt;160&gt; NUMBER OF SEQ ID NOS: 22

25 &lt;170&gt; SOFTWARE: PatentIn version 3.0

28 &lt;210&gt; SEQ ID NO: 1

30 &lt;211&gt; LENGTH: 500

32 &lt;212&gt; TYPE: DNA

34 &lt;213&gt; ORGANISM: Mycobacterium tuberculosis

37 &lt;400&gt; SEQUENCE: 1

38	catcgggtca	agcaccatga	ccggtacatc	cgtcagggtcg	tcgggacagcg	agtccagata	60
40	cggcaccggc	tggtgggttt	gctcgtcgcg	ggcgacaccg	acaaagccaa	cgtgcgcctc	120
42	cggcaaggcg	gcatgcgctt	cgtcgacatc	ccccaaaccc	gcccgaaca	caggaaccag	180
44	caggggtggc	ttggttagcc	gcgacccgac	cgtctcggcc	agcggcgtag	ggatcgggac	240
46	tggtctcgag	ggcgcatcgc	gggtggcctc	atagatcaac	agcagcgtga	gctcgcgtag	300
48	cgctgcccgg	aagccggcgt	tgctgggtgcg	ttcgtcacgc	agcgtgggtca	gtcggggccgc	360
50	ggccagtggg	tggtcaacga	catggacctg	cacggcggtg	aaccctatat	aacaatcgtg	420
52	gctcgggtccc	ctaaaagggg	gctgatacgg	gtgctgctccat	ccgcgcgacc	ggtcaacccc	480
54	gtccatatac	tcccggcatg					500

57 &lt;210&gt; SEQ ID NO: 2

59 &lt;211&gt; LENGTH: 500

61 &lt;212&gt; TYPE: DNA

63 &lt;213&gt; ORGANISM: Mycobacterium bovis

66 &lt;400&gt; SEQUENCE: 2

67	catcgggtca	agcaccatga	ccggtacatc	cgtcagggtcg	tcgggacagcg	agtccagata	60
69	cggcaccggc	tggtgggttt	gctcgtcgcg	ggcgacaccg	acaaagccaa	cgtgcgcctc	120
71	cggcaaggcg	gcatgcgctt	cgtcgacatc	ccccaaaccc	gcccgaaca	caggaaccag	180
73	caggggtggc	ttggttagcc	gcgacccgac	cgtctcggcc	agcggcgtag	ggatcgggac	240
75	tggtctcgag	ggcgcatcgc	gggtggcctc	atagatcaac	agcagcgtga	gctcgcgtag	300
77	cgctgcccgg	aagccggcgt	tgctgggtgcg	ttcgtcacgc	agcgtgggtca	gtcggggccgc	360
79	ggccagtggg	tggtcaacga	catggacctg	cacggcggtg	aaccctatat	aacaatcgtg	420
81	gctcgggtccc	ctaaaagggg	gctgatacgg	gtgctgctccat	ccgcgcgacc	ggtcaacccc	480
83	gtccatatac	tcccggcatg					500

86 &lt;210&gt; SEQ ID NO: 3

88 &lt;211&gt; LENGTH: 500

90 &lt;212&gt; TYPE: DNA

92 &lt;213&gt; ORGANISM: Mycobacterium avium

95 &lt;400&gt; SEQUENCE: 3

## RAW SEQUENCE LISTING

DATE: 02/15/2006

PATENT APPLICATION: US/10/530,983

TIME: 10:09:23

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02152006\J530983.raw

```

96 gccggtggcc agcatcgggt ccagcaccat caccgacctg ccggccagct tgtcgggcag      60
98 cgccgcaaaa tacgggaccg gccggtgggt ttgctcgtcg cgggcgacgc cgacgaagcc      120
100 cacctcggcc tcgggcagcg cggcctgcgc cgggtcgacc atgcccagcc cggcgcgagc      180
102 caccggaacc agcagcggcg gattgaccag tcgggtcccg gcggccgcgg ccaccggggt      240
104 gcggatccgg accgacttgc gcggcgcgtc gcggctggcc tcgtagacca gcatcagcgt      300
106 cagatcgcgc agcgcggccc gaaatccggc ggtgtcgggt cgttcgtcgc gcagcaccgt      360
108 cagccggggc gcggccaacg ggtggtcgat cagcacacg tccatctggt cgaggggtata      420
110 taacgatcgg gcaaagcccc gctgacacgc ttgcccgccg gccggaaacg cttaccgcc      480
112 gttcgtatac tccgggcgtg                                     500
115 <210> SEQ ID NO: 4
117 <211> LENGTH: 500
119 <212> TYPE: DNA
121 <213> ORGANISM: Mycobacterium marinum
124 <400> SEQUENCE: 4
125 cgccggtggc cagcatcggg tccagcacca ttacgggtag cccggacaag tcgtcgggca      60
127 gcgattcaag atatgggacg ggctgggtgg tctgctcatt gcgggcgata ccgacaaagc      120
129 cgaccgcgcg ctccggcagc gccgcgtgcg cttcgtcgac catgcccagt ccggcgcgca      180
131 gcaccggaac cagaagcggg gggttggcca gtcttaggcc tgtcgtggcc gcaagcgggtg      240
133 tacggatagc gacggattcg gtgggcgctg cgcgggtcgc ctcatagacc agtaccagtg      300
135 tcagctcacg caaggccttg cggaagcagc cgtttccggt gcgttcgtca cgcagcgcg      360
137 tcaggcggcg cgccgccagc ggggatcaa tgacgtggac ttccacatgg gtgaccctat      420
139 ataacaatcg gattcaagcc gctgacacgc tccccctcct cgcggcgcgg aggcgcagcc      480
141 gcccatatac tccgggcgtg                                     500
144 <210> SEQ ID NO: 5
146 <211> LENGTH: 43
148 <212> TYPE: PRT
150 <213> ORGANISM: Mycobacterium tuberculosis
154 <400> SEQUENCE: 5
156 Met Leu Arg Gly Ile Gln Ala Leu Ser Arg Pro Leu Thr Arg Val Tyr
157 1          5          10          15
159 Arg Ala Leu Ala Val Ile Gly Val Leu Ala Ala Ser Leu Leu Ala Ser
160          20          25          30
162 Trp Val Gly Ala Val Pro Gln Val Gly Leu Ala
163          35          40
165 <210> SEQ ID NO: 6
167 <211> LENGTH: 43
169 <212> TYPE: PRT
171 <213> ORGANISM: Mycobacterium bovis
174 <400> SEQUENCE: 6
176 Met Leu Arg Gly Ile Gln Ala Leu Ser Arg Pro Leu Thr Arg Val Tyr
177 1          5          10          15
179 Arg Ala Leu Ala Val Ile Gly Val Leu Ala Ala Ser Leu Leu Ala Ser
180          20          25          30
182 Trp Val Gly Ala Val Pro Gln Val Gly Leu Ala
183          35          40
185 <210> SEQ ID NO: 7
187 <211> LENGTH: 42
189 <212> TYPE: PRT
191 <213> ORGANISM: Mycobacterium avium

```

## RAW SEQUENCE LISTING

DATE: 02/15/2006

PATENT APPLICATION: US/10/530,983

TIME: 10:09:23

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02152006\J530983.raw

```

194 <400> SEQUENCE: 7
196 Met Ser Arg Glu Asn Arg Ser Arg Arg Arg Leu Ile Gly Gly Ala Tyr
197 1          5          10          15
199 Arg Ser Leu Arg Leu Leu Gly Ala Val Ala Ala Val Ala Leu Ala Ala
200          20          25          30
202 Ser Pro Leu Thr Pro Arg Thr Ser Leu Ala
203          35          40
205 <210> SEQ ID NO: 8
207 <211> LENGTH: 40
209 <212> TYPE: PRT
211 <213> ORGANISM: Mycobacterium marinum
214 <400> SEQUENCE: 8
216 Met Cys Gly Leu Lys Gln Arg Phe Thr Ser Thr Phe Arg Ala Leu Ala
217 1          5          10          15
219 Val Leu Gly Ala Val Ala Val Ser Leu Pro Ala His Gly Ser Asp Ala
220          20          25          30
222 Pro Pro Arg Ile Asp Leu Thr Ala
223          35          40
225 <210> SEQ ID NO: 9
227 <211> LENGTH: 900
229 <212> TYPE: DNA
231 <213> ORGANISM: Mycobacterium tuberculosis
234 <220> FEATURE:
236 <221> NAME/KEY: CDS
238 <222> LOCATION: (1)..(900)
241 <400> SEQUENCE: 9
242 atg ctc cgc gga atc cag gct ctc agc cgg ccc ctg acc agg gta tac      48
243 Met Leu Arg Gly Ile Gln Ala Leu Ser Arg Pro Leu Thr Arg Val Tyr
244 1          5          10          15
246 cgt gcc ttg gcg gtg atc ggt gtc ctg gca gca tcg ttg ctg gcc tca      96
247 Arg Ala Leu Ala Val Ile Gly Val Leu Ala Ala Ser Leu Leu Ala Ser
248          20          25          30
250 tgg gtc ggc gct gtc cca caa gtg ggt ctg gca gcg agt gcc ctg ccg      144
251 Trp Val Gly Ala Val Pro Gln Val Gly Leu Ala Ala Ser Ala Leu Pro
252          35          40          45
254 acc ttc gcg cac gtg gtc atc gtg gtg gag gag aac cgc tcg cag gcc      192
255 Thr Phe Ala His Val Val Ile Val Val Glu Glu Asn Arg Ser Gln Ala
256          50          55          60
258 gcc atc atc ggt aac aag tcg gct ccc ttc atc aat tcg ctg gcc gcc      240
259 Ala Ile Ile Gly Asn Lys Ser Ala Pro Phe Ile Asn Ser Leu Ala Ala
260 65          70          75          80
262 aac ggc gcg atg atg gcc cag gcg ttc gcc gaa aca cac ccg agc gaa      288
263 Asn Gly Ala Met Met Ala Gln Ala Phe Ala Glu Thr His Pro Ser Glu
264          85          90          95
266 ccg aac tac ctg gca ctg ttc gct ggc aac aca ttc ggg ttg acg aag      336
267 Pro Asn Tyr Leu Ala Leu Phe Ala Gly Asn Thr Phe Gly Leu Thr Lys
268          100          105          110
270 aac acc tgc ccc gtc aac ggc ggc gcg ctg ccc aac ctg ggt tct gag      384
271 Asn Thr Cys Pro Val Asn Gly Gly Ala Leu Pro Asn Leu Gly Ser Glu

```

## RAW SEQUENCE LISTING

DATE: 02/15/2006

PATENT APPLICATION: US/10/530,983

TIME: 10:09:23

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02152006\J530983.raw

```

272          115          120          125
274 ttg ctc agc gcc ggt tac aca ttc atg ggg ttc gcc gaa gac ttg cct      432
275 Leu Leu Ser Ala Gly Tyr Thr Phe Met Gly Phe Ala Glu Asp Leu Pro
276          130          135          140
278 gcg gtc ggc tcc acg gtg tgc agt gcg ggc aaa tac gca cgc aaa cac      480
279 Ala Val Gly Ser Thr Val Cys Ser Ala Gly Lys Tyr Ala Arg Lys His
280 145          150          155          160
282 gtg ccg tgg gtc aac ttc agt aac gtg ccg acg aca ctg tcg gtg ccg      528
283 Val Pro Trp Val Asn Phe Ser Asn Val Pro Thr Thr Leu Ser Val Pro
284          165          170          175
286 ttt tcg gca ttt ccg aag ccg cag aat tac ccc ggc ctg ccg acg gtg      576
287 Phe Ser Ala Phe Pro Lys Pro Gln Asn Tyr Pro Gly Leu Pro Thr Val
288          180          185          190
290 tcg ttt gtc atc cct aac gcc gac aac gac atg cac gac ggc tcg atc      624
291 Ser Phe Val Ile Pro Asn Ala Asp Asn Asp Met His Asp Gly Ser Ile
292          195          200          205
294 gcc caa ggc gac gcc tgg ctg aac cgc cac ctg tcg gca tat gcc aac      672
295 Ala Gln Gly Asp Ala Trp Leu Asn Arg His Leu Ser Ala Tyr Ala Asn
296          210          215          220
298 tgg gcc aag aca aac aac agc ctg ctc gtt gtg acc tgg gac gaa gac      720
299 Trp Ala Lys Thr Asn Asn Ser Leu Leu Val Val Thr Trp Asp Glu Asp
300 225          230          235          240
302 gac ggc agc agc cgc aat cag atc ccg acg gtg ttc tac ggc gcg cac      768
303 Asp Gly Ser Ser Arg Asn Gln Ile Pro Thr Val Phe Tyr Gly Ala His
304          245          250          255
306 gtg cgg ccc gga act tac aac gag acc atc agc cac tac aac gtg ctg      816
307 Val Arg Pro Gly Thr Tyr Asn Glu Thr Ile Ser His Tyr Asn Val Leu
308          260          265          270
310 tcc aca ttg gag cag atc tac gga ctg ccc aag acg ggt tat gcg acc      864
311 Ser Thr Leu Glu Gln Ile Tyr Gly Leu Pro Lys Thr Gly Tyr Ala Thr
312          275          280          285
314 aat gct ccg cca ata acc gat att tgg ggc gac tag      900
315 Asn Ala Pro Pro Ile Thr Asp Ile Trp Gly Asp
316          290          295
319 <210> SEQ ID NO: 10
321 <211> LENGTH: 299
323 <212> TYPE: PRT
325 <213> ORGANISM: Mycobacterium tuberculosis
329 <400> SEQUENCE: 10
331 Met Leu Arg Gly Ile Gln Ala Leu Ser Arg Pro Leu Thr Arg Val Tyr
332 1          5          10          15
335 Arg Ala Leu Ala Val Ile Gly Val Leu Ala Ala Ser Leu Leu Ala Ser
336          20          25          30
339 Trp Val Gly Ala Val Pro Gln Val Gly Leu Ala Ala Ser Ala Leu Pro
340          35          40          45
343 Thr Phe Ala His Val Val Ile Val Val Glu Glu Asn Arg Ser Gln Ala
344          50          55          60
347 Ala Ile Ile Gly Asn Lys Ser Ala Pro Phe Ile Asn Ser Leu Ala Ala
348 65          70          75          80

```

## RAW SEQUENCE LISTING

DATE: 02/15/2006

PATENT APPLICATION: US/10/530,983

TIME: 10:09:23

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02152006\J530983.raw

```

351 Asn Gly Ala Met Met Ala Gln Ala Phe Ala Glu Thr His Pro Ser Glu
352      85      90      95
355 Pro Asn Tyr Leu Ala Leu Phe Ala Gly Asn Thr Phe Gly Leu Thr Lys
356      100     105     110
359 Asn Thr Cys Pro Val Asn Gly Gly Ala Leu Pro Asn Leu Gly Ser Glu
360      115     120     125
363 Leu Leu Ser Ala Gly Tyr Thr Phe Met Gly Phe Ala Glu Asp Leu Pro
364      130     135     140
367 Ala Val Gly Ser Thr Val Cys Ser Ala Gly Lys Tyr Ala Arg Lys His
368 145      150     155     160
371 Val Pro Trp Val Asn Phe Ser Asn Val Pro Thr Thr Leu Ser Val Pro
372      165     170     175
375 Phe Ser Ala Phe Pro Lys Pro Gln Asn Tyr Pro Gly Leu Pro Thr Val
376      180     185     190
379 Ser Phe Val Ile Pro Asn Ala Asp Asn Asp Met His Asp Gly Ser Ile
380      195     200     205
383 Ala Gln Gly Asp Ala Trp Leu Asn Arg His Leu Ser Ala Tyr Ala Asn
384      210     215     220
387 Trp Ala Lys Thr Asn Asn Ser Leu Leu Val Val Thr Trp Asp Glu Asp
388 225      230     235     240
391 Asp Gly Ser Ser Arg Asn Gln Ile Pro Thr Val Phe Tyr Gly Ala His
392      245     250     255
395 Val Arg Pro Gly Thr Tyr Asn Glu Thr Ile Ser His Tyr Asn Val Leu
396      260     265     270
399 Ser Thr Leu Glu Gln Ile Tyr Gly Leu Pro Lys Thr Gly Tyr Ala Thr
400      275     280     285
403 Asn Ala Pro Pro Ile Thr Asp Ile Trp Gly Asp
404      290     295
407 <210> SEQ ID NO: 11
409 <211> LENGTH: 900
411 <212> TYPE: DNA
413 <213> ORGANISM: Mycobacterium bovis
416 <220> FEATURE:
418 <221> NAME/KEY: CDS
420 <222> LOCATION: (1)..(900)
423 <400> SEQUENCE: 11
424 atg ctc cgc gga atc cag gct ctc agc cgg ccc ctg acc agg gta tac      48
425 Met Leu Arg Gly Ile Gln Ala Leu Ser Arg Pro Leu Thr Arg Val Tyr
426 1      5      10      15
428 cgt gcc ttg gcg gtg atc ggt gtc ctg gca gca tcg ttg ctg gcc tca      96
429 Arg Ala Leu Ala Val Ile Gly Val Leu Ala Ala Ser Leu Leu Ala Ser
430      20      25      30
432 tgg gtc ggc gct gtc cca caa gtg ggt ctg gca gcg agt gcc ctg ccg      144
433 Trp Val Gly Ala Val Pro Gln Val Gly Leu Ala Ala Ser Ala Leu Pro
434      35      40      45
436 acc ttc gcg cac gtg gtc atc gtg gtg gag gag aac cgc tcg cag gcc      192
437 Thr Phe Ala His Val Val Ile Val Val Glu Glu Asn Arg Ser Gln Ala
438      50      55      60
440 gcc atc atc ggt aac aag tcg gct ccc ttc atc aat tcg ctg gcc gcc      240

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/530,983

DATE: 02/15/2006

TIME: 10:09:24

Input Set : A:\PTO.RJ.txt

Output Set: N:\CRF4\02152006\J530983.raw

L:14 M:270 C: Current Application Number differs, Replaced Application Number

L:16 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:972 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:17